

Europäisches Patentamt  
European Patent Office  
Office européen des brevets



(11) **EP 0 849 736 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:  
24.06.1998 Bulletin 1998/26

(51) Int Cl.<sup>6</sup>: **G11B 27/031, H04N 7/01**

(21) Application number: **97310321.1**

(22) Date of filing: **19.12.1997**

(84) Designated Contracting States:  
**AT BE CH DE DK ES FI FR GB GR IE IT LI LU MC  
NL PT SE**  
Designated Extension States:  
**AL LT LV MK RO SI**

(72) Inventors:  
• Galt, John L.  
Glendale, CA 91206 (US)  
• Pearman, James B.  
Glendale, CA 91203 (US)

(30) Priority: **20.12.1996 US 770405**

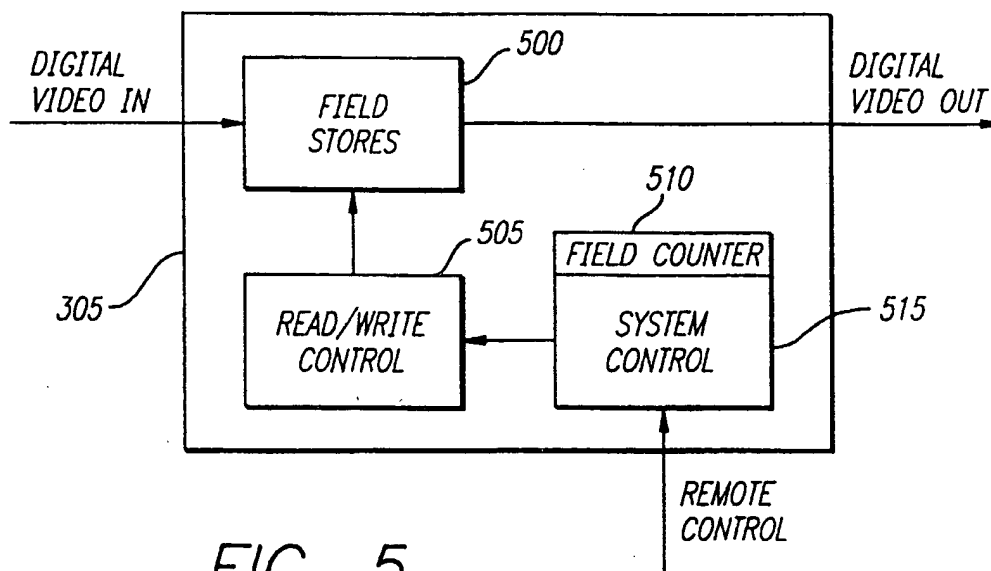
(74) Representative: **Cotter, Ivan John et al  
D. YOUNG & CO.  
21 New Fetter Lane  
London EC4A 1DA (GB)**

(71) Applicants:  
• **SONY CORPORATION**  
Tokyo 141 (JP)  
• **SONY ELECTRONICS INC.**  
Park Ridge, New Jersey 07656-8003 (US)

(54) **Method and apparatus for converting video field sequences**

(57) A resequencer (305) changes the ABCD grouping of a sequence of video fields derived from a motion picture file via 3-2 pulldown so as to allow that sequence of video fields to be cut into or edited with a second sequence of similarly derived video fields, the cut or edit occurring at any desired position, without any artifacts.

The resequencer comprises eight field stores (500). In each video field time, an input video field is stored into one field store and an output video field is read from a different field store. The particular field stores into which fields are read and written in each video field time are chosen so as to achieve the desired change in ABCD grouping.



**FIG. 5**

**EP 0 849 736 A1**

## Description

This invention relates to converting video field sequences. The invention can for instance be implemented in the form of methods of and apparatus for changing phase of video material taken from 24 frame per second film. The invention can be applied to video post-production editing systems, and in particular to editing systems which are used to edit program material taken from standard 24 frame per second film.

Movies and videotapes give the illusion of motion by playing back rapidly a sequence of still images called "frames." Commercial movies are designed to be played back at 24 frames per second. In contrast, under the American NTSC television standard, videotapes are designed to be played back at 29.97 frames per second. Under high-definition video standard SMPTE-240M, high-definition videotapes are designed to be played back at 30 frames per second. In Europe, under the PAL and SECAM television standards, videotapes are designed to be played back at 25 frames per second.

Besides the difference in frame rates, video playback differs in another important way from movie playback. In video, the screen image is divided up into a large number of horizontal scan lines. The American NTSC standard provides for 525 scan lines on the screen (not all of which are visible), whereas the European PAL standard provides for 625 scan lines. In videotapes, the scan lines that correspond to each frame are divided into two "fields," the first containing the even-numbered scan lines and the second the odd-numbered scan lines. On videotape, to record a frame, one records first one of the fields and then the other field. Thus, a video frame contains two fields, and an SMPTE-240M videotape is designed to be played back at 60 fields per second, corresponding to 30 frames per second.

It is common to move program material from motion picture film to videotape. A machine which can perform this function is called a "telecine." For further background on telecines, see U.S. Patent No. 5,428,387, issued to the present inventors. Nowadays, it is also common to move program material from videotape to film, particularly inasmuch as high definition video equipment allows this operation to be carried out without degradation in image quality. Either process requires a way of bridging the gap between the 24 frames per second used in film and the 25, 29.97, or 30 frames per second used in videotape.

When converting 24 frame per second film to 29.97/30 frame per second videotape, the standard process used to bridge the frame rate gap is called "3-2 pulldown." In this process, even-numbered frames of the film are recorded as two successive fields, which are identical (except that they correspond to different sets of scanlines). Odd-numbered frames of the film are recorded as three successive fields, which are also identical (except that they correspond to different sets of scanlines). Thus, every pair of frames in the film, which

takes 1/12th of a second to play back at the film playback rate of 24 frames per second, is converted into five video fields, which also take 1/12th of a second to play back at a video playback rate of 60 fields per second.

As depicted in Fig. 1, videotapes derived from a film 100 by means of 3-2 pulldown consist of a sequence of groups of fields such as 125, 130, 135 and 140, each field group corresponding to one film frame such as 105. There are actually four different types of field groups, which we call A, B, C, and D. An A field group consists of two fields corresponding to one videotape frame and to one film frame. A B field group consists of three fields, the first two belonging to one videotape frame and the third to the next videotape frame, with the three fields corresponding to one film frame. A C field group consists of two fields, corresponding to two different videotape frames but only one film frame. A D field group consists of three videotape fields, the first belonging to one videotape frame and the other two belonging to the next videotape frame, but all corresponding to one film frame.

It is important to note the following difference between A and C field groups. Even though both consist of two fields corresponding to one film frame, the first field of an A field group is the first field of a video frame, and thus it corresponds to even-numbered scanlines. In contrast, the first field of a C field group is the second field of a video frame, and thus it corresponds to odd-numbered scanlines. There is a similar difference between B and D field groups. The first field of a B field group is the first field of a video frame, and thus it corresponds to even-numbered scanlines. In contrast, the first field of a D field group is the second field of a video frame, and thus it corresponds to odd-numbered scanlines.

A videotape derived from film by means of 3-2 pulldown thus consists of a sequence of field groups in the order ABCD, ABCD, ABCD, etc. Notably, the first two fields of the videotape will be an A field group, and thus it will be possible to know, just from the location of a field within the videotape, whether it belongs to an A, B, C, or D field group. In modern videotape editing, the frames are referred to by "time codes," which take the form hh:mm:ss:ff, where hh denotes hours, mm minutes, ss seconds, and ff frames (not fields). Because there are six sets of A, B, C, and D field groups in each second's worth of videotape, each set of field groups encompassing five video frames, it is possible to know just from the video frame number ff of the time code whether the fields of that video frame belong to A, B, C, or D field groups. In particular, the first video frame of each second consists of two fields forming an A field group, the second video frame consists of two fields that make up part of a B field group, the third video frame consists of one field from a B field group and one from a C field group, and so forth.

It is common, in the process of editing videotape, to cut, wipe, or dissolve between two or more sources of video. Post-production video editing systems of the

types known in the art permit this cutting, wiping, or dissolving to be performed at any video frame boundary.

Consider a typical editing scenario depicted in **Fig. 2**. The user of the editing system has two source videotapes, tape X and tape Y, derived from film by means of 3-2 pulldown. The user wishes to copy the first six video frames, numbered 0 through 5 in **Fig. 2**, from tape X onto the output, followed by five video frames, numbered 0 through 4, from tape Y. Because both tapes were obtained by 3-2 pulldown, tape X's video frames 0 through 5 come from five film frames and have an ABCDA sequence, while tape X's video frames 0 through 4 come from four film frames and have an ABCD sequence. If this operation were carried out on a prior art editing system, first the six video frames 0 through 5 in tape X will be copied to the output, which will give the output tape an initial ABCDA sequence. Then the five video frames 0 through 4 from tape Y will be copied to the output. This additional copy operation will result in the output having an undesirable irregular ABCDAAB-CD sequence, as shown in **Fig. 2**.

Suppose now that the operation depicted in **Fig. 2** were a dissolve rather than a cut, so that contents of the two tapes are blended, with an increasing proportion of tape Y content being used as time advances. It is readily seen that the output video frame numbered 7 consists of two fields. The first field of output frame 7 is derived from tape X's B film frame (spanning the video frames numbered 6 and 7) blended with tape Y's B film frame (spanning video frames numbered 1 and 2). The second field of output frame 7, however, is derived from a different film frame on tape X, the C film frame spanning video frames 7 and 8. This is also undesirable.

As this example shows, even if the source videotapes X and Y are derived from film via 3-2 pulldown and have perfect ABCD sequences, an editing operation can potentially result in irregularities in the ABCD sequence of the output videotape. These irregularities can give rise to visible artifacts when the output videotape is converted back to film. Even a single cut can result in all the videotape after the cut having an offset ABCD sequence, so that the time code no longer allows one to determine whether the fields of a frame belong to A, B, C, or D field groups.

One technique for avoiding this difficulty is to limit the resolution at which cutting, wiping and dissolving take place so that these operations are only allowed at the boundary of an ABCD structure. Since an ABCD structure is five videotape frames long, i.e., one-sixth of a second, this greatly restricts the editor's ability to cut as artistic considerations dictate (e.g., to synchronize with sound effects). This loss of creative control is unacceptable in practice.

Respective different aspects of the invention are set forth in claims 1, 8 and 14 hereof.

A preferred form of implementation of the invention described below creates an editing system for videotape which allows cut and dissolve at the highest temporal

resolution consistent with preserving the ABCD sequence of field groups. This is achieved by means of a modest alteration of existing videotape editing systems, without requiring any radical changes in existing systems.

The preferred form of implementation of the invention makes use of a novel apparatus called a resequencer. The resequencer contains eight field stores, i.e., devices capable of storing the contents of a video field at the resolution corresponding to the video equipment being used (e.g., NTSC or SMPTE-240M). A video tape player is connected to the input of the resequencer, and the resequencer generates video at its output. One resequencer is used for each source of video going into the editing system, so that for example if the editing system is capable of combining video from three video tape players to produce the output videotape three resequencers are needed.

The preferred form of implementation of the invention also makes use of a controller interfaced to the resequencers and commanding them. The controller commands the resequencer to perform the appropriate transformations between its input and output so that the resequencers' outputs have an ABCD sequence of field groups which matches the desired ABCD sequence in the output videotape.

Through the combined action of the controller and the resequencer, it is thus possible to achieve the objective of performing cuts, wipes, and dissolves between two video sources while preserving the output videotape's regular ABCD sequence of field groups. These cuts, wipes and dissolves can be carried out with the maximum, temporal resolution which the sources permit, so that the cut, wipe, or dissolve can start with the beginning of any A, B, C or D field group in each source and be placed at the beginning of any A, B, C or D field group in the output.

The invention will now be further described, by way of illustrative and nonlimiting example, with reference to the accompanying drawings, in which:

**Figure 1** depicts four types of videotape fields A, B, C and D, which exist in a 60 field per second videotape derived from motion picture film;

**Figure 2** depicts the video field sequences in an editing operation which, under prior art editing systems, would result in an irregular ABCD sequence on the output videotape;

**Figure 3** depicts how a resequencer embodying the invention fits into an overall video editing system;

**Figure 4** shows how the video editing operation of **Fig. 2** would be carried out using the resequencer;

**Figure 5** depicts components which make up the resequencer; and

**Figure 6 to 9** depict frame stores which are used when the resequencer is operating in ABCD, BCDA, CDAB, and DABC modes, respectively.

In the following description, for purposes of explanation, specific nomenclature is set forth to provide a thorough understanding of an embodiment of the present invention. However, it will be apparent to one skilled in the art that these specific details are not required in order to practice the present invention. In other instances, well known circuits and devices are shown in block diagram form in order not to obscure the disclosure.

The embodiment of the invention is designed to be used as part of a video post production editing system of the kind depicted schematically in Fig. 3. Such a video editing system employs a controller 330 which allows user control of the editing process. The video editing system further comprises two or more source videotape players 300, 320, two or more resequencers 305, 325 embodying the invention, a switcher 310, and a videotape recorder 315 on which the output video is recorded. One resequencer is connected between each source videotape player and the switcher 310. Video editing systems also typically contain a console, a video monitor, an audio mixer, speakers, and other equipment not shown.

Fig. 4 shows how the editing operation depicted in Fig. 2 works when the resequencers embodying the invention are in use. The first six video frames on tape X are written to the output as before, as five field groups in the sequence ABCDA. However, the five starting video frames from tape Y are transformed from ABCD into BCDA by the resequencers before being written onto the output tape. This transformation, also referred to as a "phase change," ensures that the output tape has the desired regular ABCD sequence of field groups.

The transformation of an ABCD sequence of field groups into a BCDA sequence requires the resequencer to duplicate certain of the video fields and drop others. Specifically, for the initial A field group from tape Y in Fig. 4, comprising two fields, the resequencer must generate the three fields corresponding to the B field group which is to be written onto the output tape. The resequencer must do so by duplicating the first field in the input A field group to obtain both the first and third fields of the output B field group. In addition, the B field group from tape Y in Fig. 4 needs to be converted to a C field group, and thus its three fields need to be transformed into two. This transformation is accomplished by dropping the last of the three fields of the B field group and reversing the order of the other two. The reversal is required because the first field in a B field group is the first field of a video frame and thus corresponds to even scanlines, while the first field of a C field group is the second field of a video frame and thus corresponds to odd scanlines. Similarly, the C field group from tape Y needs to be converted to a D field group. This conver-

sion is accomplished just as the A-to-B conversion was, by duplicating the first field in the input C field group to obtain both the first and third fields of the output D field group. Finally, the D field group from tape Y needs to be converted into an A field group; again, this is accomplished by dropping the last field of the D field group and reversing the order of the other two.

While the example of Fig. 4 required the resequencer to convert the sequence of field groups on tape from ABCD to BCDA, other editing operations may require that a sequence of field groups be converted from ABCD to CDAB or from ABCD to DABC, or that the output sequence be ABCD just like the input sequence. It is thus seen that a resequencer must operate in four different modes, which may conveniently be designated as ABCD (meaning that an ABCD input is converted to ABCD output), BCDA (meaning that an ABCD input is converted to BCDA output), CDAB (meaning that an ABCD input is converted to CDAB output), and DABC (meaning that an ABCD input is converted to DABC output).

The resequencer 305 or 325 internally has the structure shown in Fig. 5. It consists of a set of eight field stores 500, a read-write control 505 for the field stores 500, and a system control 515 which communicates with the edit controller 330. The incoming and outgoing fields are preferably communicated in digital form, in any of a number of digital video formats well known in the art. In each field time interval (1/60th of a second), the incoming field is written into one of the eight field stores 500, while another of the field stores 500 is read and directed to the resequencer's output. The resequencer also contains a field counter 510 which increments each time a new field is output, i.e. every 60th of a second. The number stored in this counter is referred to as the field count. The field count is taken modulo 10 so that it varies only from 0 to 9; this may conveniently be accomplished by resetting it to 0 each time it would otherwise increment past 9. The system control 515 decides which field stores to write and read in each field time interval as a function of the mode and the field count. The manner in which this is decided is explained below.

Figs. 6 to 9 define, for each of the four modes, which of the eight field stores 500 are read and written in each field time interval as a function of the field count and the mode. Fig. 6 shows which field stores are read and written as a function of the field count when the resequencer is operating in ABCD mode. When the field count is 0, the first column of Fig. 6 shows that field store 1 is written and field store 7 is read. When the field count is 1, the second column shows that field store 2 is written and field store 8 is read. When the field count is 2, field store 3 is written and field store 1 is read. When the field count is 3, field store 4 is written and field store 2 is read. When the field count is 4, field store 3 is read and no field store is written. When the field count is 5, field store 6 is written and field store 4 is read. When the field count is 6, field store 5 is written and field store 3 is read. When

the field count is 7, field store 8 is written and field store 6 is read. When the field count is 8, field store 7 is written and field store 5 is read. Finally, when the field count is 9, field store 8 is read and no field store is written.

It should be noted from Fig. 6 that the resequencer introduces a time delay between its input and its output. Specifically, in ABCD mode, the incoming field that arrives when the field count is 0 is sent to the output when the field count is 2, i.e. two field times (1/30th of a second) later. The edit controller 330 takes this time delay into account in the overall synchronization of the system.

Fig. 7 similarly defines which field stores are read and written when the resequencer is operating in BCDA mode. Fig. 7 shows, for example, that when the field count is 0, field store 1 is written and field store 7 is read. When the field count is 1, field store 2 is written and field store 7 is read. When the field count is 2, field store 3 is written and field store 1 is read. The field stores which are read and written for the remaining values of field count may easily be read from the remaining columns of Fig. 7, as was done in the discussion of Fig. 6 above. Fig. 8 shows the field stores which are read and written when the resequencer is operating in CDAB mode, and Fig. 9 shows the field stores which are read and written when the resequencer is operating in DABC mode. When the resequencer is operating in these latter two modes, the time delay between input and output is 3 field times (1/20th of a second).

In the specification of the resequencer's operation which has been given so far, the resequencer has always worked with a sequence of field groups on the source videotape which starts with an A field group. However, the resequencer is also capable of starting with a B, C, or D field group. To start with a B, C, or D field group, the resequencer's field count starts with a value different from zero. For example, suppose that the resequencer operates in BCDA mode, as in Fig. 7, and the sequence of field groups in the source starts with a C field group. It is then merely necessary to start the resequencing process with a field count of 5 rather than 0, as is readily seen by examining Fig. 7. The controller 330 preferably sends the resequencer an indication of what type of field group the first source field on each source tape belongs to. This indication lets the resequencer know which field count to start out with.

The choice of which field stores to read and write in Figs. 6 to 9 is only exemplary. Those skilled in the art will recognize that many other choices are possible which still achieve the necessary ABCD sequence conversions called for by the chosen mode.

Although each resequencer 305, 325 is shown in Fig. 3 as a separate box, it will be clear to those skilled in the art that a number of resequencers can be packaged inside one box, sharing a common power supply and other elements, such as the RS-422 link to the edit controller 330. Furthermore, although the preferred embodiment has been described in terms of videotape ed-

iting, it will be appreciated by those skilled in the art that the invention can also be practiced with video which is stored on other media, such as laser discs.

## Claims

1. A method for converting an input sequence of video fields comprising a first plurality of A, B, C, and D field groups to an output sequence of video fields comprising a second, different plurality of A, B, C, and D field groups, said method comprising the steps of

receiving an input video field,  
selecting a first field store from a plurality of field stores,  
writing said input video field into said first field store,  
selecting a second field store from said plurality of field stores,  
reading an output video field from said second field store, and  
outputting said output video field,

said steps being performed repeatedly so that each field group in said output sequence consists of video fields belonging to exactly one field group in said input sequence.

2. The method of claim 1, further comprising the step of receiving an indication of whether said input sequence of video fields begins with an A, B, C, or D field group.

3. The method of claim 1, further comprising the step of modifying a field count, where said step of selecting a second field store takes as input said field count.

4. The method of claim 3, where said step of modifying said field count comprises the substeps of incrementing said field count by 1 and setting said field count to 0 if said incremented value equals 10.

5. The method of claim 1, further comprising the step of receiving a mode indication, where said step of selecting a second field store takes as an input said mode indication.

6. The method of claim 5, where said mode indication takes on at least four distinct values.

7. The method of claim 1, where at least one A field group in said output sequence corresponds to a B, C, or D field group in said input sequence.

8. An apparatus for converting an input sequence of

video fields comprising a first plurality of A, B, C, and D field groups to an output sequence of video fields comprising a second, different plurality of A, B, C, and D field groups, said apparatus comprising

a plurality of field stores,  
 means for receiving an input video field,  
 means for selecting a first field store from said plurality of field stores,  
 means for writing said input video field into said first field store,  
 means for selecting a second field store from said plurality of field stores,  
 means for reading an output video field from said second field store, and  
 means for outputting said output video field,

so that each field group in said output sequence consists of video fields belonging to exactly one field group in said input sequence.

9. The apparatus of claim 8, further comprising means for receiving an indication of whether said input sequence of video fields begins with an A, B, C, or D field group.
10. The apparatus of claim 8, further comprising means for modifying a field count, where said means for selecting a first field store receives as input said field count.
11. The apparatus of claim 10, where said means for modifying said field count comprises means for incrementing said field count by 1 and means for setting said field count to 0 if said incremented value equals 10.
12. The apparatus of claim 8, further comprising means for receiving a mode indication, where said means for selecting a second field store receives as input said mode indication.
13. The apparatus of claim 8, where at least one A field group in said output sequence corresponds to an B, C, or D field group in said input sequence.
14. A video post-production editing system comprising

a videotape player generating a sequence of video fields,  
 a resequencing device coupled to the output of said videotape player for converting said sequence of video fields to an output sequence of video fields, said resequencing device comprising a plurality of field stores,

means for receiving a video field from said sequence of video fields,

means for selecting a first field store from said plurality of field stores,  
 means for writing said input video field into said first field store,  
 means for selecting a second field store from said plurality of field stores,  
 means for reading an output video field from said second field store, and  
 means for outputting said output video field,

a switcher coupled to the output of said resequencing device for switching among a plurality of sources of program material, one of said plurality of sources being said resequencing device,  
 a videotape recorder coupled to the output of said switcher, and  
 an edit controller coupled to and controlling said videotape player, resequencing device, switcher, and videotape recorder.

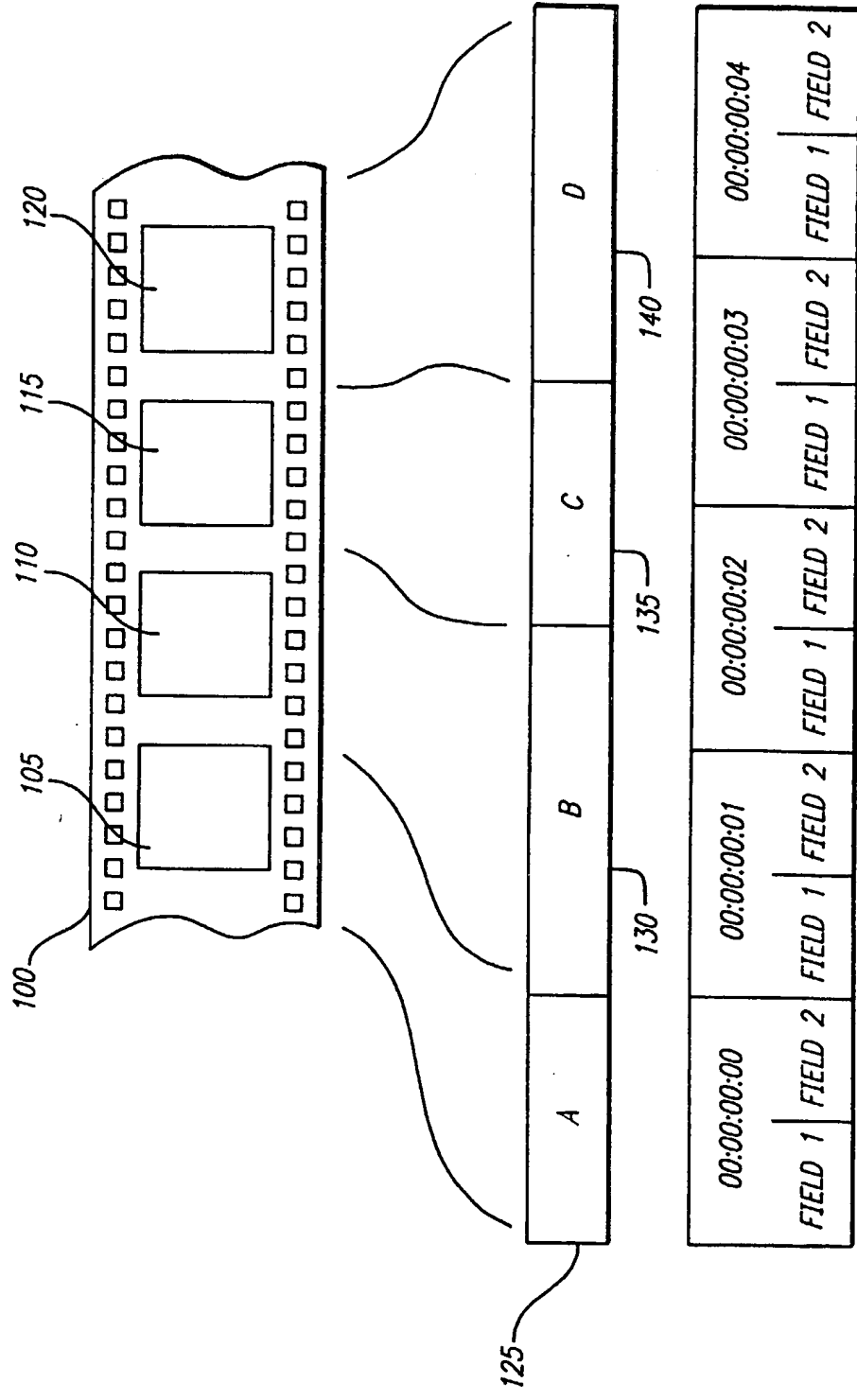


FIG. 1  
PRIOR ART

FIG. 2 *PRIOR ART*

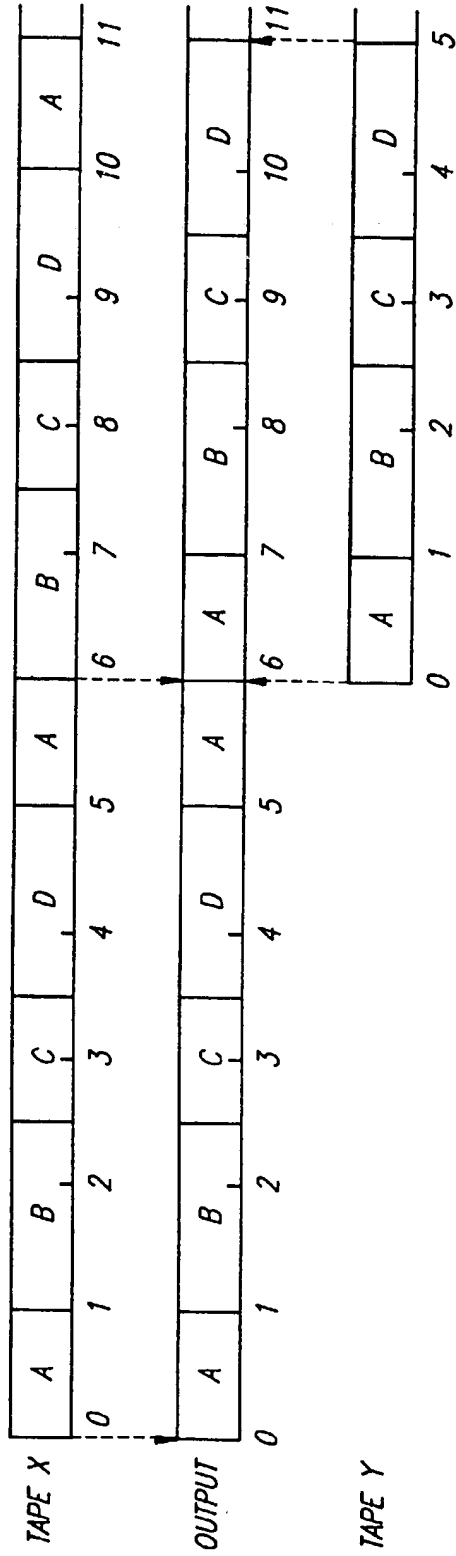
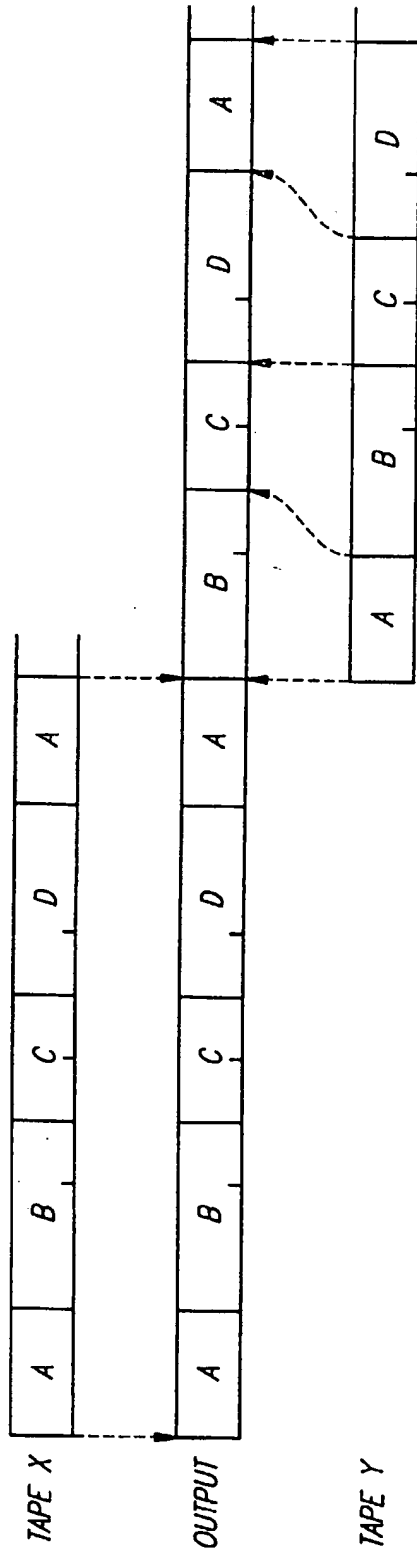


FIG. 4





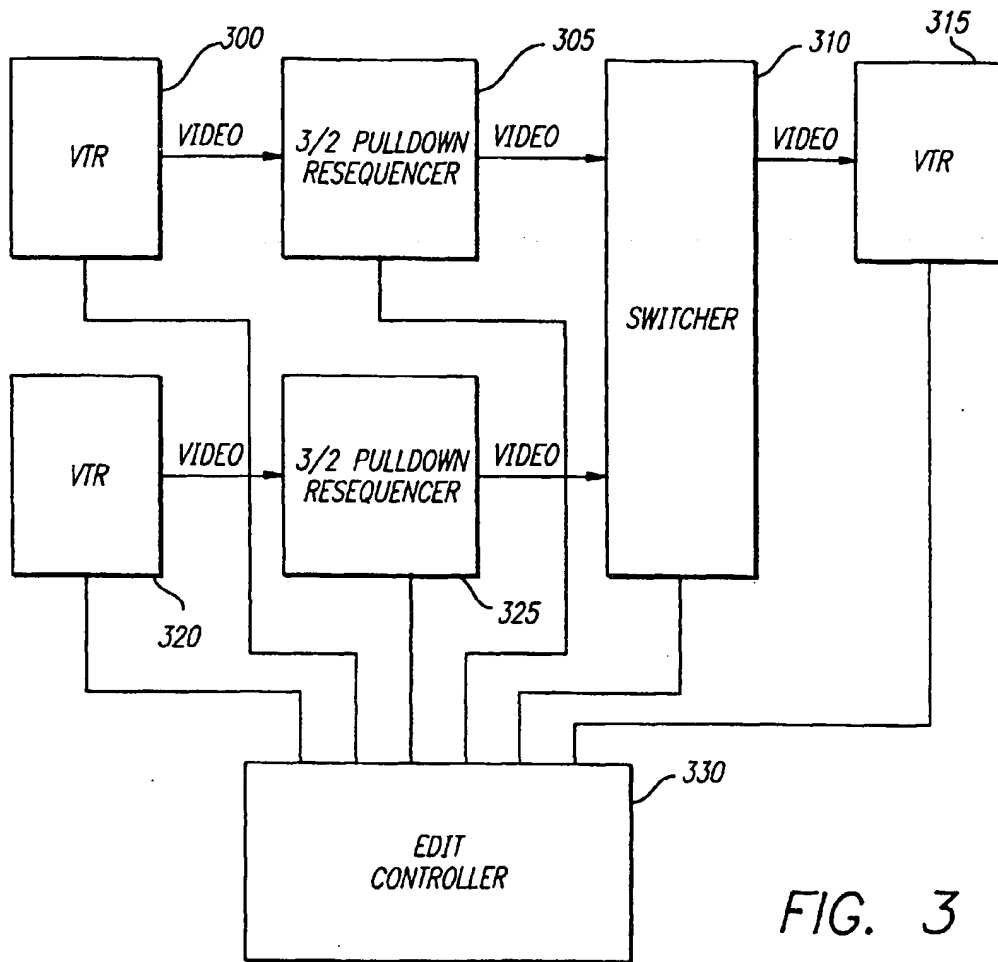


FIG. 3

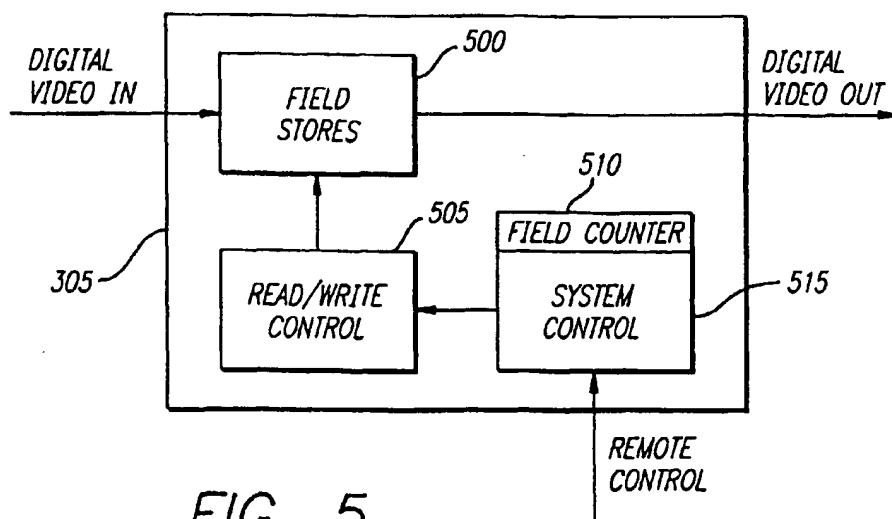


FIG. 5

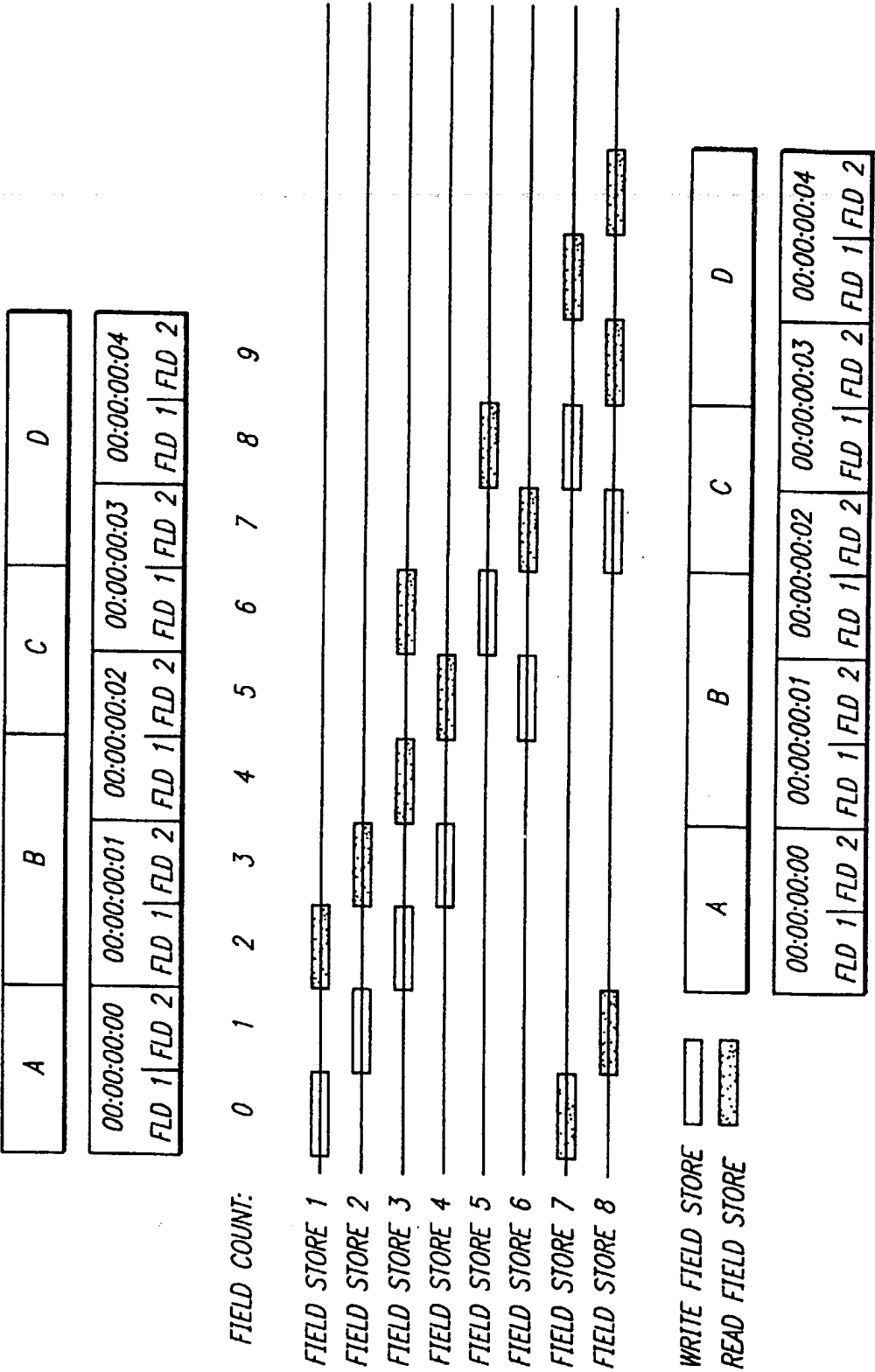


FIG. 6

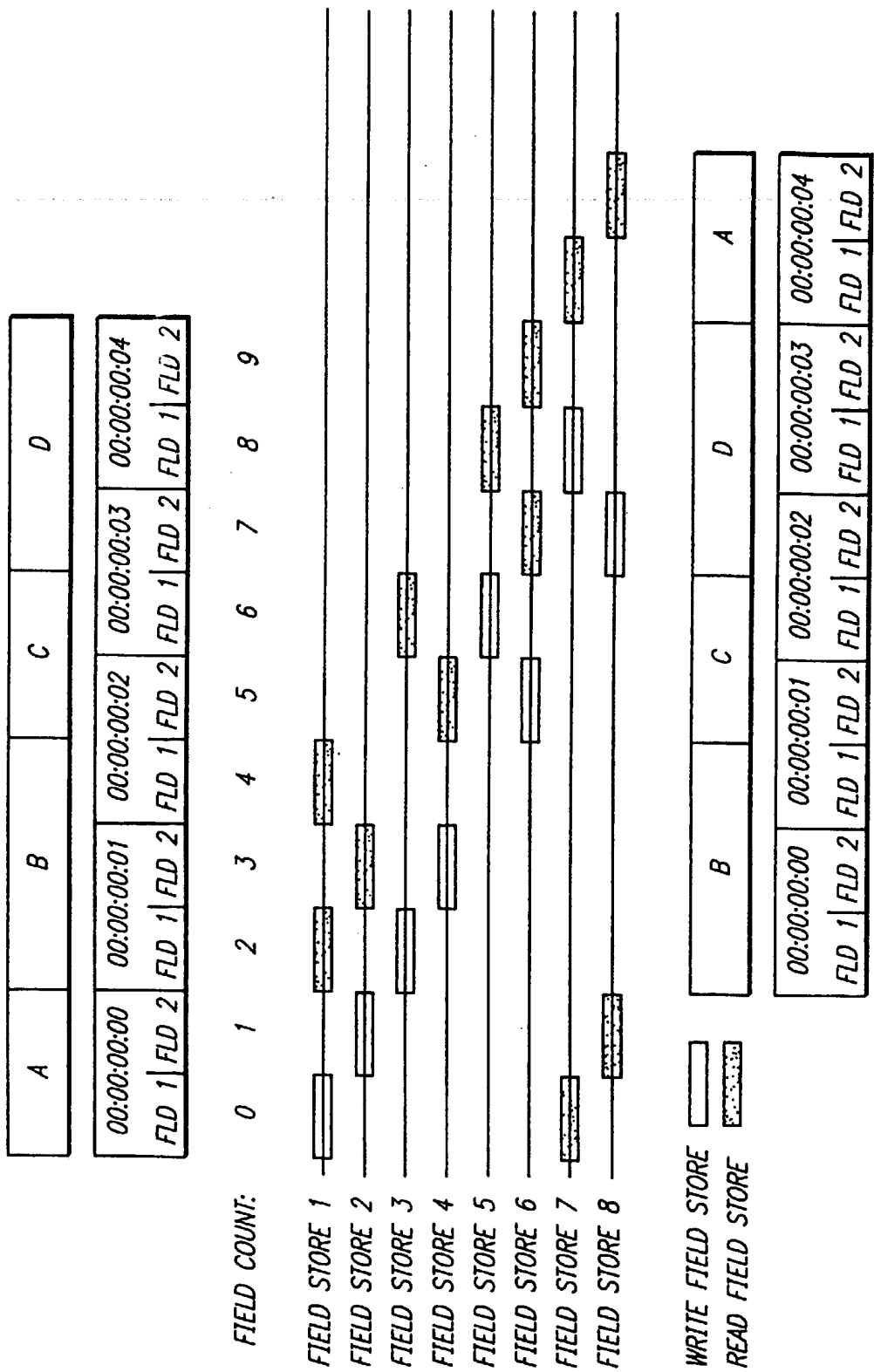


FIG. 7

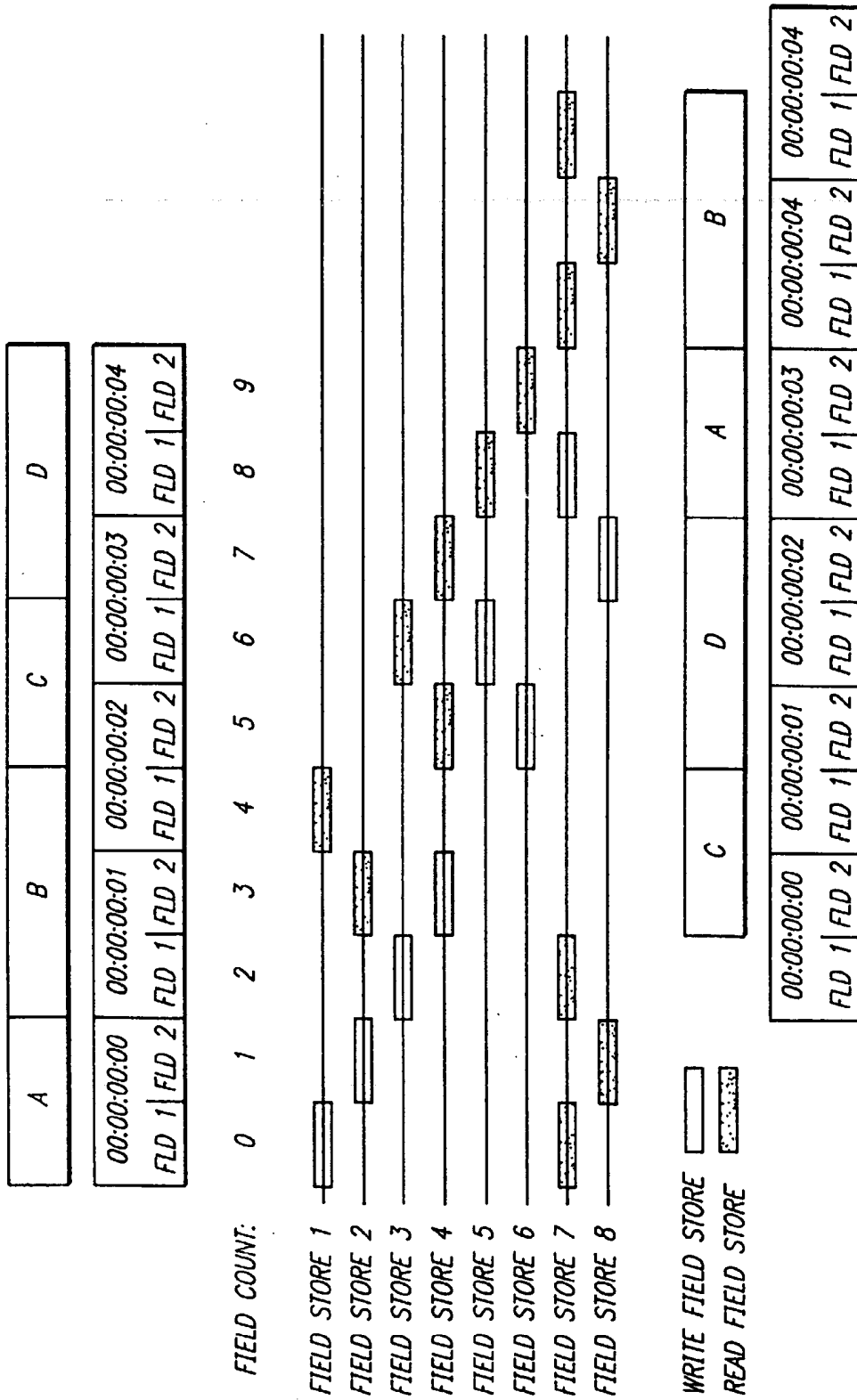


FIG. 8

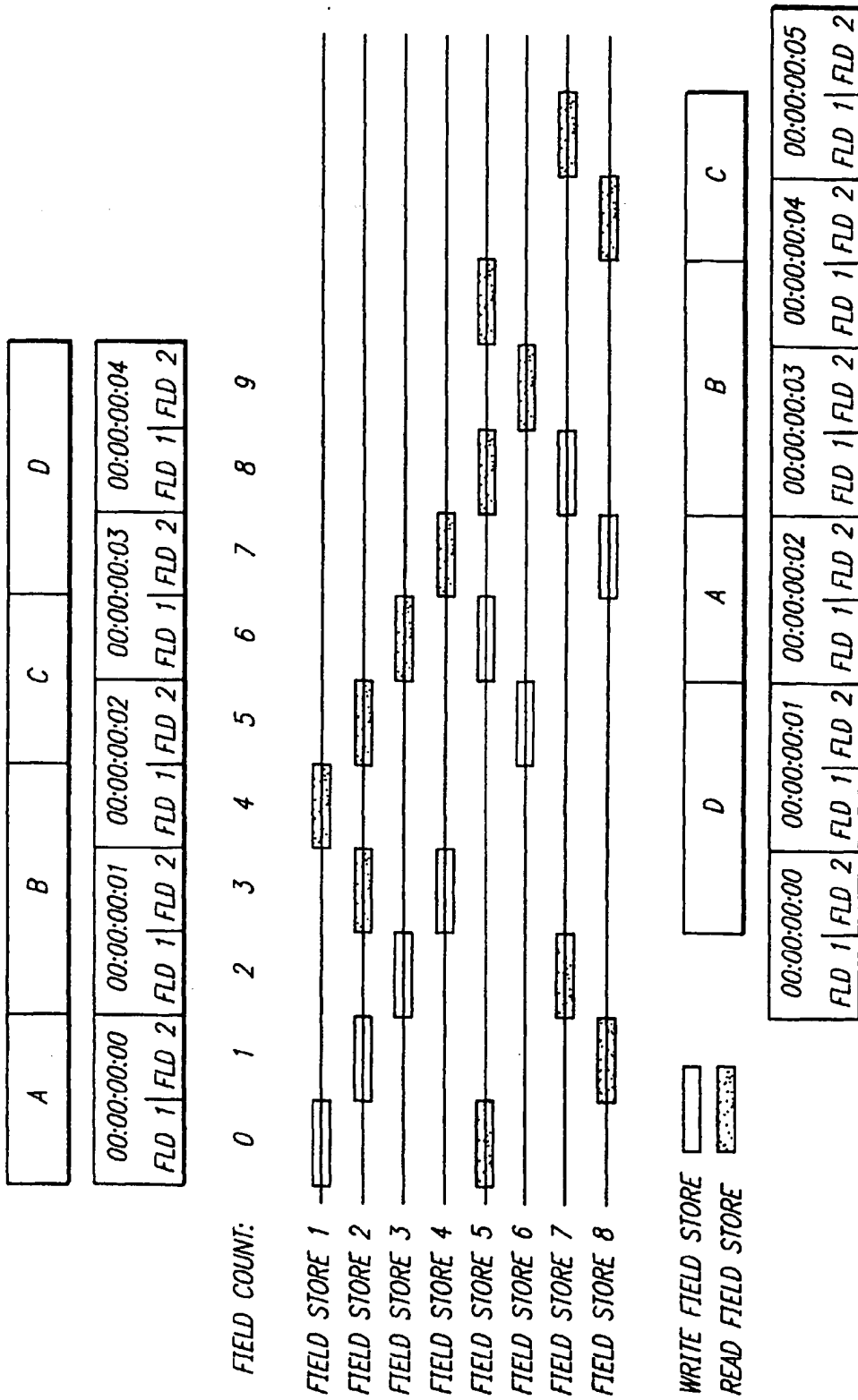


FIG. 9



European Patent  
Office

# EUROPEAN SEARCH REPORT

Application Number  
EP 97 31 0321

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.6)
X	US 4 998 167 A (JAQUA DOUGLAS A)	1,2,5-9, 12,13	G11B27/031 H04N7/01
A	* column 9, line 25 - column 15, line 30; figure 9 *	3,10,14	
A	WO 94 01971 A (AVID TECHNOLOGY INC) * page 8, line 5 - page 11, line 10 *	1,8,14	
A	GB 2 265 782 A (SNELL & WILCOX LTD) * page 3, line 7 - page 7, line 15 *	1,8,14	
A	EP 0 685 968 A (VICTOR COMPANY OF JAPAN) * column 16, line 25 - column 17, line 45 *	1,8,14	
			TECHNICAL FIELDS SEARCHED (Int.Cl.6)
			G11B H04N
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 23 March 1998	Examiner Beaudoin, O
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons &amp; : member of the same patent family, corresponding document</p>			

EPO FORM 1503 03/82 (P4/C01)